

OIKE

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/802,755A

TIME: 14:18:05

Input Set : A:\2555-001.ST25.txt

Output Set: N:\CRF3\08012001\I802755A.raw

ENTERED

3 <110> APPLICANT: Darst, Seth A
 4 Campbell, Elizabeth A.
 6 <120> TITLE OF INVENTION: CRYSTAL OF BACTERIA CORE RNA POLYMERASE WITH RIFAMPICIN
 8 <130> FILE REFERENCE: 2555-1-001
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/802,755A
 11 <141> CURRENT FILING DATE: 2001-03-09
 13 <160> NUMBER OF SEQ ID NOS: 4
 15 <170> SOFTWARE: PatentIn version 3.1
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 1525
 19 <212> TYPE: PRT
 20 <213> ORGANISM: Thermus aquaticus
 22 <220> FEATURE:
 23 <221> NAME/KEY: MISC_FEATURE
 24 <222> LOCATION: (1247)..(1247)
 25 <223> OTHER INFORMATION: Xaa can be any amino acid.
 28 <400> SEQUENCE: 1
 30 Met Lys Lys Glu Val Arg Lys Val Arg Ile Ala Leu Ala Ser Pro Glu
 31 1 5 10 15
 34 Lys Ile Arg Ser Trp Ser Tyr Gly Glu Val Glu Lys Pro Glu Thr Ile
 35 20 25 30
 38 Asn Tyr Arg Thr Leu Lys Pro Glu Arg Asp Gly Leu Phe Asp Glu Arg
 39 35 40 45
 42 Ile Phe Gly Pro Ile Lys Asp Tyr Glu Cys Ala Cys Gly Lys Tyr Lys
 43 50 55 60
 46 Arg Gln Arg Phe Glu Gly Lys Val Cys Glu Arg Cys Gly Val Glu Val
 47 65 70 75 80
 50 Thr Arg Ser Ile Val Arg Arg Tyr Arg Met Gly His Ile Glu Leu Ala
 51 85 90 95
 54 Thr Pro Ala Ala His Ile Trp Phe Val Lys Asp Val Pro Ser Lys Ile
 55 100 105 110
 58 Gly Thr Leu Leu Asp Leu Phe Ala Thr Glu Leu Glu Gln Val Leu Tyr
 59 115 120 125
 62 Phe Asn Lys Tyr Ile Val Leu Asp Pro Lys Gly Ala Val Leu Asp Gly
 63 130 135 140
 66 Val Pro Val Glu Lys Arg Gln Leu Leu Thr Asp Glu Glu Tyr Arg Glu
 67 145 150 155 160
 70 Leu Arg Tyr Gly Lys Gln Glu Thr Tyr Pro Leu Pro Ala Gly Val Asp
 71 165 170 175
 74 Ala Leu Val Lys Asp Gly Glu Glu Val Val Lys Gly Gln Glu Leu Ala
 75 180 185 190
 78 Pro Gly Val Val Ser Arg Met Asp Gly Val Gly Ser Leu Pro Leu Pro
 79 195 200 205
 82 Arg Arg Val Arg Val Asp Tyr Leu Arg Lys Glu Arg Ala Ala Leu Arg
 83 210 215 220
 86 Ile Pro Leu Ser Ala Trp Val Glu Lys Glu Pro Tyr Arg Pro Gly Glu
 87 225 230 235 240

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/802,755A

TIME: 14:18:05

Input Set : A:\2555-001.ST25.txt

Output Set: N:\CRF3\08012001\I802755A.raw

```

90 Val Leu Ala Glu Leu Ser Glu Pro Tyr Leu Phe Arg Ala Glu Glu Ser
91                245                250                255
94 Gly Val Val Glu Leu Lys Asp Leu Ala Glu Gly His Leu Ile Tyr Leu
95                260                265                270
98 Arg Gln Glu Glu Glu Val Val Ala Arg Tyr Phe Leu Pro Ala Gly Met
99                275                280                285
102 Thr Pro Leu Val Val Glu Gly Glu Ile Val Glu Val Gly Gln Pro Leu
103                290                295                300
106 Ala Glu Gly Lys Gly Leu Leu Arg Leu Pro Arg His Met Thr Ala Lys
107 305                310                315                320
110 Glu Val Glu Ala Glu Glu Glu Gly Asp Ser Val His Leu Thr Leu Phe
111                325                330                335
114 Leu Glu Trp Thr Glu Pro Lys Asp Tyr Lys Val Ala Pro His Met Asn
115                340                345                350
118 Val Ile Val Pro Glu Gly Ala Lys Val Gln Ala Gly Glu Lys Ile Val
119                355                360                365
122 Ala Ala Ile Asp Pro Glu Glu Glu Val Ile Ala Gln Ala Glu Gly Val
123                370                375                380
126 Val His Leu His Glu Pro Ala Ser Ile Leu Val Val Lys Ala Arg Val
127 385                390                395                400
130 Tyr Pro Phe Glu Asp Asp Val Glu Val Thr Thr Gly Asp Arg Val Ala
131                405                410                415
134 Pro Gly Asp Val Leu Ala Asp Gly Gly Lys Val Lys Ser Glu Ile Tyr
135                420                425                430
138 Gly Arg Val Glu Val Asp Leu Val Arg Asn Val Val Arg Val Val Glu
139                435                440                445
142 Ser Tyr Asp Ile Asp Ala Arg Met Gly Ala Glu Ala Ile Gln Glu Leu
143                450                455                460
146 Leu Lys Glu Leu Asp Leu Glu Lys Leu Glu Arg Glu Leu Leu Glu Glu
147 465                470                475                480
150 Met Lys His Pro Ser Arg Ala Arg Arg Ala Lys Ala Arg Lys Arg Leu
151                485                490                495
154 Glu Val Val Arg Ala Phe Leu Asp Ser Gly Asn Arg Pro Glu Trp Met
155                500                505                510
158 Ile Leu Glu Ala Val Pro Val Leu Pro Pro Asp Leu Arg Pro Met Val
159                515                520                525
162 Gln Val Asp Gly Gly Arg Phe Ala Thr Ser Asp Leu Asn Asp Leu Tyr
163                530                535                540
166 Arg Arg Leu Ile Asn Arg Asn Asn Arg Leu Lys Lys Leu Leu Ala Gln
167 545                550                555                560
170 Gly Ala Pro Glu Ile Ile Ile Arg Asn Glu Lys Arg Met Leu Gln Glu
171                565                570                575
174 Ala Val Asp Ala Val Ile Asp Asn Gly Arg Arg Gly Ser Pro Val Thr
175                580                585                590
178 Asn Pro Gly Ser Glu Arg Pro Leu Arg Ser Leu Thr Asp Ile Leu Ser
179                595                600                605
182 Gly Lys Gln Gly Arg Phe Arg Gln Asn Leu Leu Gly Lys Arg Val Asp
183                610                615                620
186 Tyr Ser Gly Arg Ser Val Ile Val Val Gly Pro Gln Leu Lys Leu His

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/802,755A

TIME: 14:18:05

Input Set : A:\2555-001.ST25.txt

Output Set: N:\CRF3\08012001\I802755A.raw

```

187 625          630          635          640
190 Gln Cys Gly Leu Pro Lys Arg Met Ala Leu Glu Leu Phe Lys Pro Phe
191          645          650          655
194 Leu Leu Lys Lys Met Glu Glu Lys Ala Phe Ala Pro Asn Val Lys Ala
195          660          665          670
198 Ala Arg Arg Met Leu Glu Arg Gln Arg Asp Ile Lys Asp Glu Val Trp
199          675          680          685
202 Asp Ala Leu Glu Glu Val Ile His Gly Lys Val Val Leu Leu Asn Arg
203          690          695          700
206 Ala Pro Thr Leu His Arg Leu Gly Ile Gln Ala Phe Gln Pro Val Leu
207 705          710          715          720
210 Val Glu Gly Gln Ser Ile Gln Leu His Pro Leu Val Cys Glu Ala Phe
211          725          730          735
214 Asn Ala Asp Phe Asp Gly Asp Gln Met Ala Val His Val Pro Leu Ser
215          740          745          750
218 Ser Phe Ala Gln Ala Glu Ala Arg Ile Gln Met Leu Ser Ala His Asn
219          755          760          765
222 Leu Leu Ser Pro Ala Ser Gly Glu Pro Leu Ala Lys Pro Ser Arg Asp
223          770          775          780
226 Ile Ile Leu Gly Leu Tyr Tyr Ile Thr Gln Val Arg Lys Glu Lys Lys
227 785          790          795          800
230 Gly Ala Gly Met Ala Phe Ala Thr Pro Glu Glu Ala Leu Ala Ala Tyr
231          805          810          815
234 Glu Arg Gly Glu Val Ala Leu Asn Ala Pro Ile Val Val Ala Gly Arg
235          820          825          830
238 Glu Thr Ser Val Gly Arg Leu Lys Phe Val Phe Ala Asn Pro Asp Glu
239          835          840          845
242 Ala Leu Leu Ala Val Ala His Gly Leu Leu Asp Leu Gln Asp Val Val
243          850          855          860
246 Thr Val Arg Tyr Leu Gly Arg Arg Leu Glu Thr Asn Pro Gly Arg Ile
247 865          870          875          880
250 Leu Phe Ala Arg Ile Val Gly Glu Ala Val Gly Asp Glu Lys Val Ala
251          885          890          895
254 Gln Glu Leu Ile Gln Met Asp Val Pro Gln Glu Lys Asn Ser Leu Lys
255          900          905          910
258 Asp Leu Val Tyr Gln Ala Phe Leu Arg Leu Gly Met Glu Lys Thr Ala
259          915          920          925
262 Arg Leu Leu Asp Ala Leu Lys Tyr Tyr Gly Phe Thr Leu Ser Thr Thr
263          930          935          940
266 Ser Gly Ile Ile Thr Ile Gly Ile Asp Asp Ala Val Ile Pro Glu Glu
267 945          950          955          960
270 Lys Gln Arg Tyr Leu Glu Glu Ala Asp Arg Lys Leu Arg Gln Ile Glu
271          965          970          975
274 Gln Ala Tyr Glu Met Gly Phe Leu Thr Asp Arg Glu Arg Tyr Asp Gln
275          980          985          990
278 Val Ile Gln Leu Trp Thr Glu Thr Thr Glu Lys Val Thr Gln Ala Val
279          995          1000          1005
282 Phe Asn Asn Phe Glu Glu Asn Tyr Pro Phe Asn Pro Leu Tyr Val
283          1010          1015          1020

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/802,755A

TIME: 14:18:05

Input Set : A:\2555-001.ST25.txt

Output Set: N:\CRF3\08012001\I802755A.raw

286 Met Ala Gln Ser Gly Ala Arg Gly Asn Pro Gln Gln Ile Arg Gln
 287 1025 1030 1035
 290 Leu Cys Gly Met Arg Gly Leu Met Gln Lys Pro Ser Gly Glu Thr
 291 1040 1045 1050
 294 Phe Glu Val Pro Val Arg Ser Ser Phe Arg Glu Gly Leu Thr Val
 295 1055 1060 1065
 298 Leu Glu Tyr Phe Ile Ser Ser His Gly Ala Arg Lys Gly Gly Ala
 299 1070 1075 1080
 302 Asp Thr Ala Leu Arg Thr Ala Asp Ser Gly Tyr Leu Thr Arg Lys
 303 1085 1090 1095
 306 Leu Val Asp Val Ala His Glu Ile Val Val Arg Glu Ala Asp Cys
 307 1100 1105 1110
 310 Gly Thr Thr Lys Tyr Ile Ser Val Pro Leu Phe Gln Met Asp Glu
 311 1115 1120 1125
 314 Val Thr Arg Thr Leu Arg Leu Arg Lys Arg Ser Asp Ile Glu Ser
 315 1130 1135 1140
 318 Gly Leu Tyr Gly Arg Val Leu Ala Arg Glu Val Glu Ala Leu Gly
 319 1145 1150 1155
 322 Arg Arg Leu Glu Glu Gly Arg Tyr Leu Ser Leu Glu Asp Val His
 323 1160 1165 1170
 326 Phe Leu Ile Lys Ala Ala Glu Ala Gly Glu Val Arg Glu Val Pro
 327 1175 1180 1185
 330 Val Arg Ser Pro Leu Thr Cys Gln Thr Arg Tyr Gly Val Cys Gln
 331 1190 1195 1200
 334 Lys Cys Tyr Gly Tyr Asp Leu Ser Met Ala Arg Pro Val Ser Ile
 335 1205 1210 1215
 338 Gly Glu Ala Val Gly Val Val Ala Ala Glu Ser Ile Gly Glu Pro
 339 1220 1225 1230
 342 Gly Thr Gln Leu Thr Met Arg Thr Phe His Thr Gly Gly Xaa Ala
 343 1235 1240 1245
 346 Val Gly Thr Asp Ile Thr Gln Gly Leu Pro Arg Val Ile Glu Leu
 347 1250 1255 1260
 350 Phe Glu Ala Arg Arg Pro Lys Ala Lys Ala Val Ile Ser Glu Ile
 351 1265 1270 1275
 354 Asp Gly Val Val Arg Ile Glu Glu Gly Glu Asp Arg Leu Ser Val
 355 1280 1285 1290
 358 Phe Val Glu Ser Glu Gly Phe Ser Lys Glu Tyr Lys Leu Pro Lys
 359 1295 1300 1305
 362 Asp Ala Arg Leu Leu Val Lys Asp Gly Asp Tyr Val Glu Ala Gly
 363 1310 1315 1320
 366 Gln Pro Leu Thr Arg Gly Ala Ile Asp Pro His Gln Leu Leu Glu
 367 1325 1330 1335
 370 Ala Lys Gly Pro Glu Ala Val Glu Arg Tyr Leu Val Asp Glu Ile
 371 1340 1345 1350
 374 Gln Lys Val Tyr Arg Ala Gln Gly Val Lys Leu His Asp Lys His
 375 1355 1360 1365
 378 Ile Glu Ile Val Val Arg Gln Met Leu Lys Tyr Val Glu Val Thr
 379 1370 1375 1380
 382 Asp Pro Gly Asp Ser Pro Leu Leu Glu Gly Gln Val Leu Glu Lys

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/802,755A

TIME: 14:18:05

Input Set : A:\2555-001.ST25.txt

Output Set: N:\CRF3\08012001\I802755A.raw

```

383      1385      1390      1395
386 Trp Asp Val Glu Ala Leu Asn Glu Arg Leu Ile Ala Glu Gly Lys
387      1400      1405      1410
390 Val Pro Val Ala Trp Lys Pro Leu Leu Met Gly Val Thr Lys Ser
391      1415      1420      1425
394 Ala Leu Ser Thr Lys Ser Trp Leu Ser Ala Ala Ser Phe Gln Asn
395      1430      1435      1440
398 Thr Thr His Val Leu Thr Glu Ala Ala Ile Ala Gly Lys Lys Asp
399      1445      1450      1455
402 Glu Leu Ile Gly Leu Lys Glu Asn Val Ile Leu Gly Arg Leu Ile
403      1460      1465      1470
406 Pro Ala Gly Thr Gly Ser Asp Phe Val Arg Phe Thr Gln Val Val
407      1475      1480      1485
410 Asp Gln Arg Thr Leu Lys Ala Ile Glu Glu Ala Arg Lys Glu Ala
411      1490      1495      1500
414 Val Glu Ala Lys Glu Lys Glu Ala Pro Arg Arg Pro Val Arg Arg
415      1505      1510      1515
418 Glu Gln Pro Gly Lys Gly Leu
419      1520      1525
422 <210> SEQ ID NO: 2
423 <211> LENGTH: 1119
424 <212> TYPE: PRT
425 <213> ORGANISM: Thermus aquaticus
427 <220> FEATURE:
428 <221> NAME/KEY: MISC_FEATURE
429 <222> LOCATION: (695)..(696)
430 <223> OTHER INFORMATION: Xaa can be any amino acid
433 <220> FEATURE:
434 <221> NAME/KEY: MISC_FEATURE
435 <222> LOCATION: (1247)..(1247)
436 <223> OTHER INFORMATION: Xaa can be any amino acid
439 <400> SEQUENCE: 2
441 Met Lys Ile Lys Arg Phe Gly Arg Ile Arg Glu Val Ile Pro Leu Pro
442 1      5      10      15
445 Pro Leu Thr Glu Ile Gln Val Glu Ser Tyr Lys Lys Ala Leu Gln Ala
446      20      25      30
449 Asp Val Pro Pro Glu Lys Arg Glu Asn Val Gly Ile Gln Ala Ala Phe
450      35      40      45
453 Lys Glu Thr Phe Pro Ile Glu Gly Asp Lys Gly Lys Gly Gly Leu
454      50      55      60
457 Val Leu Asp Phe Leu Glu Tyr Arg Ile Gly Asp Pro Pro Phe Ser Gln
458 65      70      75      80
461 Asp Glu Cys Arg Glu Lys Asp Leu Thr Tyr Gln Ala Pro Leu Tyr Ala
462      85      90      95
465 Arg Leu Gln Leu Ile His Lys Asp Thr Gly Leu Ile Lys Glu Asp Glu
466      100     105     110
469 Val Phe Leu Gly His Leu Pro Leu Met Thr Glu Asp Gly Ser Phe Ile
470      115     120     125
473 Ile Asn Gly Ala Asp Arg Val Ile Val Ser Gln Ile His Arg Ser Pro

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/802,755A

DATE: 08/01/2001

TIME: 14:18:06

Input Set : A:\2555-001.ST25.txt

Output Set: N:\CRF3\08012001\I802755A.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number
L:342 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:613 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2